# Session F: Differentiation by Finite Differences

## F.1 Introduction

In this exercise, we extend our use of finite differences to obtain numerical solutions of derivatives. After developing the necessary algebra, we will look at some simple functions where the solutions are analytic. This will enable us to test the accuracy of the approximate numerical solutions that we obtain. However, it should go without saying that the techniques that we will develop in this session are applicable to a wide range of problems, where the solution will not always be so obvious.

## F.2 The differential operator

We met the differential operator, $D$, very briefly in Session D. We defined the operator as:

$$D\{f(m)\} = (df/dx)_m$$

where the derivative is evaluated at the point $x = a_0 + mh$. In order to develop a numerical approach to this kind of operation, we must find a finite difference representation of the operator, $D$. We shall follow a similar approach to that taken in the last session where we dealt with the problem of interpolation. Once again we assume that we have the values of a function, $y$, on a mesh of *evenly* spaced $x$ values, the spacing being given by $h$.

By using a Taylor series expansion, we can express the value of the function at $m = 1$ in terms of the value of the function at $m = 0$ in the following way:

$$y(x+h) = y(x) + h(dy/dx)_{y(x)} + (h^2/2!)(d^2y/dx^2)_{y(x)} + \cdots + (h^n/n!)(d^ny/dx^n)_{y(x)} + \ldots \quad \text{(F.1)}$$

The subscripts indicate that the derivatives are evaluated at the origin, $m = 0$. Expressing this equation in terms of the differential operator, $D$, we obtain:

$$y(x+h) = (1 + hD + (h^2/2!)D^2 + \cdots + (h^n/n!)D^n + \ldots)\{y(x)\} \quad \text{(F.2)}$$

Now for the tricky part! We can write this expression in terms of the more familiar forward shift operator, $E$, if we note that:

1. We can write $y(x+h) = E\{y(x)\}$.
2. The expansion of $e^z$ is $(1 + z + (z^2/2!) + \cdots + (z^n/n!) + \ldots)$.

Thus, we can simply express $E$ in terms of $D$, as shown below:

$$E\{y(x)\} = (e^{hD})\{y(x)\} \quad \text{(F.3)}$$

Check this for yourself. This procedure may seem a bit odd, but it is similar in practice to what we did when we derived the interpolation formulae in the last session. What we have ended up with is the relationship between the *forward shift* operator, $E$, and the *differential* operator, $D$:

$$E = e^{hD}$$

The point is that the forward shift operation is something we can handle numerically. So to express $D$ in terms of $E$, which is what we really want:

$$hD = \ln(E) \tag{F.4}$$

We can now use the table of relationships between the different types of operators (which we derived in Session D) to express the differential operator in terms of the forward and backward difference operators:

$$hD = \ln(1 + \Delta) \tag{F.5}$$

and

$$hD = -\ln(1 - \nabla) \tag{F.6}$$

If we take the forward difference form in equation F.5, we can expand the logarithm in terms of a power series. Recall that the Maclaurin series for $\ln(1 + x)$ is:

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \tag{F.7}$$

This means that we can expand the operator function in equation F.5 to give the differential operator in terms of powers of the forward difference operator:

$$hD = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \ldots \tag{F.8}$$

If we did the same with the backward differences formula, using equation F.6, we would obtain:

$$hD = \nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \ldots \tag{F.9}$$

You might like to verify this for yourself. As with our previous use of series like this one, we can truncate the expression at a desired point. For example, using the forward difference formula (F.8), the first order approximation of the derivative of a function, $f(x)$, is:

$$\begin{aligned} df/dx &= D\{f(0)\} \approx (1/h)\Delta\{f(0)\} \\ df/dx &\approx (1/h)[f(1) - f(0)] \end{aligned} \tag{F.10}$$

The validity of this approximation is, hopefully, rather self-evident. To first order, the first derivative is given by the difference between the values of two neighbouring points (evaluated at $m = 0$ and 1) divided by the step size. You didn't need two pages of algebra to tell you this! However, in equations F.8 and F.9 you now have expressions that enable you to calculate the first derivative (or gradient) of the function, at any given point, to arbitrary accuracy.

Let's put all this into practice. Then, hopefully, all will become clear.

---

## Exercise F.1

1. Make a spreadsheet for the function, $f(x) = x^5 + 4x^4 + 20x^3 + x^2 + 10x + 2$ for integer values of $x$ from 0 to 20.

2. Use equation F.10 to evaluate the differential of $f(x)$ at $x = 0, 1, 2, \ldots$.
3. Compare this with the analytic result, expressing the difference as a percentage of the analytic result. Are results for all values of $x$ equally accurate?
4. Now use equation F.8 with terms up to $\Delta^2$. Again compare with the analytic result.
5. Repeat with 3rd, 4th and 5th order differences (you will need to extend equation F.8 to the next higher term to obtain the expression to 5th order).
6. Hopefully it should be obvious why we do not need to go to 6th order differences!

---

---

### Exercise F.2

1. Make a spreadsheet for the same function at values of $x$ between 0 and 3.0 in steps of 0.1.
2. Compare the results for $x = 0$, $x = 1$ and $x = 2$ with those obtained in Exercise F.1. Notice that the accuracy compared with the analytic result is now significantly improved.
3. Repeat for $x = 0$ to 0.5 in steps of 0.01. Again note the improved accuracy.

---

## F.3 Higher order differentials

We can extend our approach to higher order differentials by using:

$$D^r\{f(m)\} = \{d^r f/dx^r\}_m$$

To simplify the notation, we will denote higher order differentials simply as $f^{(r)}$. Thus, based on equation F.8 we can write:

$$h^r f^{(r)} = h^r D^r\{f(0)\} = (\Delta - (1/2)\Delta^2 + (1/3)\Delta^3 - (1/4)\Delta^4 \ldots)^r\{f(0)\}$$

Taking the common factor of $\Delta^r$ outside the bracket gives:

$$h^r f^{(r)} = \Delta^r \left(1 - (1/2)\Delta + (1/3)\Delta^2 - (1/4)\Delta^3 \ldots\right)^r \{f(0)\}$$

Expanding based on $(1+z)^r$, where $z = [-(1/2)\Delta + (1/3)\Delta^2 - (1/4)\Delta^3 \ldots]$, gives:

$$\begin{aligned} h^r f^{(r)} = \Delta^r(1 & \\ & + r[-(1/2)\Delta + (1/3)\Delta^2 - (1/4)\Delta^3 \ldots] \\ & + (r(r-1)/2!)[-(1/2)\Delta + (1/3)\Delta^2 - (1/4)\Delta^3 \ldots]^2 \\ & \ldots)\{f(0)\} \end{aligned}$$

Collecting up the terms to the same order in $\Delta$, this simplifies down to:

$$h^r f^{(r)} = \left(\Delta^r - \frac{r}{2}\Delta^{r+1} + \frac{r(3r+5)}{24}\Delta^{r+2}\dots\right)\{f(0)\} \tag{F.11}$$

Similarly, in terms of the backwards difference operator, it can be shown that:

$$h^r f^{(r)} = \left(\nabla^r + \frac{r}{2}\nabla^{r+1} + \frac{r(3r+5)}{24}\nabla^{r+2} + \dots\right)\{f(0)\} \tag{F.12}$$

---

### Exercise F.3

1. Make a *forward differences* column for the function used in Exercise F.1. Calculate the second order differential using equation F.11 using just the first term, which will be of order $\Delta^2$ for a second order derivative.
2. Repeat using three terms, $\Delta^2$ to $\Delta^4$.

---

It should be noted that, contrary to expectation, for a given value of $r$, taking a larger number of terms in the expression for the $r^{\text{th}}$ derivative might lead to erroneous (less accurate) results. We shall discuss this point in more detail later. For the present you should note that the backward and forward differences are generally slowly convergent. However, it turns out that the convergence can be significantly improved by using central differences.

## F.4 Central differences

In Session D, we showed that the central difference operator, $\delta$, could be written in terms of the forward shift operator, $E$, such that:

$$\delta = E^{1/2} - E^{-1/2}$$

In this session we showed that the forward shift operator and the differential operator are related by $E = e^{hD}$. Thus, we can also relate the central difference operator using:

$$\delta = e^{hD/2} - e^{-hD/2}$$

Once again, it will be convenient to write this expression in a slightly different form. This time, we can use the hyperbolic sine function, given that $\sinh(x) = 0.5(e^x - e^{-x})$. Thus, an alternative way to write the central difference operator is:

$$\delta/2 = \sinh(hD/2)$$

or,

$$hD = 2\sinh^{-1}(\delta/2)$$

A convenient way to expand this is in the form $\sinh^{-1}(x)/x$.

$$hD = [(2/\delta)\sinh^{-1}(\delta/2)]\delta$$

Thus, for orders of differentation which are even (i.e. $r = 2s$ where $s$ is an integer):

$$h^{2s} f^{(2s)}(0) = h^{2s} D^{2s}\{f(0)\} = [(2/\delta)\sinh^{-1}(\delta/2)]^{2s}\delta^{2s}\{f(0)\}$$

For $s = 1$ we can show that:

$$\frac{d^2 f}{dx^2} = \frac{1}{h^2}\left[\delta^2 - \frac{\delta^4}{12} + \frac{\delta^6}{90}\dots\right]\{f(0)\} \tag{F.13}$$

You can show this for yourself if you look up the expansion of $\sinh^{-1}(x)$. I don't expect you to know all these expansions (see note at the end); I just want you to know where the formulae like equation F.13 come from.

Thus, to first order, the second derivative, expressed in terms of the central difference operator, becomes:

$$\frac{d^2 f}{dx^2} \approx \frac{1}{h^2}[\delta^2]\{f(0)\}$$
$$\frac{d^2 f}{dx^2} = \frac{1}{h^2}[\delta]\left\{f(\frac{1}{2}) - f(-\frac{1}{2})\right\}$$
$$\frac{d^2 f}{dx^2} = \frac{1}{h^2}[f(1) - 2f(0) + f(-1)]$$

---

## Exercise F.4

1. Make a central difference column for the expression in Exercise F.1. Calculate the second order differential using the first order term, $\delta^2$.
2. Repeat using terms up to $\delta^4$.
3. Contrast the results from Exercise F.3 and F.4.

---

Unfortunately, the scheme represented by equation F.13 does not work for odd orders of differentiation, that is, when $r = 2s + 1$ where $s$ is integer. This is because the expansion ends us up with odd powers of $\delta$, but this would mean we would need values of our function at intermediate mesh points $f(m/2)$ where we have no measurement. A slightly different formulation exists to handle odd orders. I won't derive it, but simply quote the result in case you ever need it, or come across it in another context:

$$h^{2s+1} f^{(2s+1)}(0) = h^{2s+1} D^{2s+1}\{f(0)\} = (1/\mu)[(2/\delta\sinh^{-1}(\delta/2)]^{2s+1}\mu\delta^{2s+1}\{f(0)\}$$

This form can be expanded in terms of the points $f(-1), f(0), f(1)$, etc.

It may occur to you that just as we have considered higher orders of differentiation as $D^n$, we can also perform integration by using $D^{-1}$. However, we shall look at integration in detail in the next session.

## F.5 Truncation errors

What we have seen in this session is that, given a differential operator $D^r$, there is no unique finite difference representation. We can use forward, backward and central differences, each with their own degree of approximation. It is therefore important to have some idea of how accurate a given representation will be. For example, is the forward difference formula better or worse than the backward or central difference form?

Let's take a look at this and consider a second order derivative. First, stand the problem on its head and try to see exactly what a particular finite difference representation actually means. We know that if we consider a function $f(x)$ at $x = a_0$ (i.e. $f(a_0 + mh)$ with $m = 0$, or written in brief, $f(0)$), then the second order forward difference operator gives us:

$$\Delta^2\{f(0)\} = f(2) - 2f(1) + f(0)$$

But now express $f(1)$ in terms of the Taylor series:

$$f(1) = f(0) + h(df/dx)_0 + (h^2/2!)(d^2f/dx^2)_0 + \cdots + (h^n/n!)(d^nf/dx^n)_0 + \ldots$$

Similarly, for $f(-1)$ we could write:

$$f(-1) = f(0) - h(df/dx)_0 + (h^2/2!)(d^2f/dx^2)_0 + \cdots + (-1)^n(h^n/n!)(d^nf/dx^n)_0 + \ldots$$

In exactly the same way we can derive $f(2)$ and $f(-2)$ as shown below:

$$\begin{aligned} f(2) &= f(0) + (2h)(df/dx)_0 + ((2h)^2/2!)(d^2f/dx^2)_0 + \cdots + (+1)^n((2h)^n/n!)(d^nf/dx^n)_0 + \ldots \\ f(-2) &= f(0) - (2h)(df/dx)_0 + ((2h)^2/2!)(d^2f/dx^2)_0 + \cdots + (-1)^n((2h)^n/n!)(d^nf/dx^n)_0 + \ldots \end{aligned}$$

Thus, if we were to approximate a second order differential with a second order forward difference form (equation F.11):

$$d^2f/dx^2 \approx \Delta^2\{f(0)\}/h^2$$

we can now expand this in terms of $f(1)$, $f(2)$, etc.

$$\Delta^2\{f(0)\}/h^2 = [f(2) - 2f(1) + f(0)]/h^2 = (d^2f/dx^2)_0 + h(d^3f/dx^3)_0 + O(h^2)$$

Here, $O(h^2)$ means other terms involving $h$ to the power of two, or higher powers. It is more meaningful to write this instead of adding the ellipsis ("…"). If we truncate the series at the first term we get:

$$\Delta^2\{f(0)\}/h^2 = (d^2f/dx^2)_0 + O(h)$$

or

$$(d^2f/dx^2)_0 = \Delta^2\{f(0)\}/h^2 - O(h)$$

and the result is good to order $h$.

By contrast, if we use central differences we find that:

$$(1/h^2)[\delta^2]\{f(0)\} = (1/h^2)[f(1) - 2f(0) + f(-1)] = d^2f/dx^2 + O(h^2)$$

and the estimate is good to second order. The central difference form is therefore more accurate.

## F.6 Partial differentials

The above treatment can be straightforwardly extended to cover partial differentiation of a function of two or more variables. We can define partial differential operators in the same way as we did for the complete differential, only now our function is evaluated at the mesh points of a two (or higher) dimensional grid:

$$\begin{aligned} D_x\{f(m,n)\} &= (\partial f/\partial x)_{m,n} \\ D_y\{f(m,n)\} &= (\partial f/\partial y)_{m,n} \\ &\cdots \end{aligned}$$

where the derivatives are evaluated at $x = a_0 + mh_x$ and $y = b_0 + nh_y$, with $h_x$ and $h_y$ being the spacing of the mesh points in the $x$ and $y$ directions respectively, and $a_0$ and $b_0$ are the initial values for $x$ and $y$ for our mesh.

We can proceed exactly as before and determine expressions for these operators in terms of the finite difference operators, $\Delta_x$, $\Delta_y$, $\nabla_x$, $\nabla_y$, $\delta_x$ and $\delta_y$, which act only on the indicated variables and leave the other variable(s) unaffected. In other words, these operators only use points along a horizontal or vertical line on our mesh, never a diagonal one. As an example, the first order approximation for $\partial f/\partial x$ using forward differences would be:

$$\begin{aligned} \partial f/\partial x = D_x\{f(0,0)\} &\approx (1/h_x)\Delta_x\{f(0,0)\} \\ &= (1/h_x)\,[f(1,0) - f(0,0)] \end{aligned}$$

Note that the second argument (the $y$ position on the mesh) is unchanged.

Similarly, the lowest order approximation for the second derivative using central differences is:

$$\partial^2 f/\partial x^2 \approx (1/h_x^2)[\delta_x^2]\{f(0,0)\} = (1/h_x^2)\,[f(1,0) - 2f(0,0) + f(-1,0)] \tag{F.14}$$

### F.6.1 Mixed derivatives

Where we have to be careful is when we wish to evaluate mixed derivatives, such as:

$$\frac{\partial^2 f}{\partial x \partial y} = D_x\left\{D_y\{f(0,0)\}\right\} = D_y\left\{D_x\{f(0,0)\}\right\}$$

In terms of first order approximations using forward differences, it is straightforward to write:

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &\approx \frac{1}{h_x h_y}\Delta_x\{\Delta_y\{f(0,0)\}\} \\ &= \frac{1}{h_x h_y}\Delta_x\{[f(0,1) - f(0,0)]\} \\ &= \frac{1}{h_x h_y}\,[f(1,1) - f(0,1) - f(1,0) + f(0,0)] \end{aligned} \tag{F.15}$$

Similarly, for backward differences:

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &\approx \frac{1}{h_x h_y}\nabla_x\{\nabla_y\{f(0,0)\}\} \\ &= \frac{1}{h_x h_y}\,[f(0,0) - f(-1,0) - f(0,-1) + f(-1,-1)] \end{aligned} \tag{F.16}$$

However, for central differences we cannot write:

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y} &\approx \frac{1}{h_x h_y}\delta_x\{\delta_y\{f(0,0)\}\} \\ &= \frac{1}{h_x h_y}\left[f(\frac{1}{2},\frac{1}{2}) - f(-\frac{1}{2},\frac{1}{2}) - f(\frac{1}{2},-\frac{1}{2}) + f(-\frac{1}{2},-\frac{1}{2})\right]\end{aligned}$$

as this would required evaluation of $f$ at intermediate mesh points.

This is an example of the different requirements for *odd* orders of differentials (here order 1). In this case, the first order approximation is given by:

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y} &\approx \frac{1}{h_x h_y}\mu_x\delta_x\{\mu_y\delta_y\{f(0,0)\}\} \\ &= \frac{1}{4h_x h_y}\left[f(1,1) - f(-1,1) - f(1,-1) + f(-1,-1)\right] \qquad \text{(F.17)}\end{aligned}$$

---

## Exercise F.5

1. Using the data from Exercise C.4, and the given parameter values for the gradient ($p_1$) and intercept ($p_0$), calculate $\chi^2(p_0, p_1)$.
2. Using step sizes of $h_{p_0} = 0.001$ and $h_{p_1} = 0.01$, calculate $\chi^2(p_0, p_1 \pm h_{p_1})$, $\chi^2(p_0 \pm h_{p_0}, p_1)$ and $\chi^2(p_0 \pm h_{p_0}, p_1 \pm h_{p_1})$ (this gives you 8 more values of $\chi^2$).
3. Use equations F.14 and F.17 to calculate central difference approximations for the second derivatives of $\chi^2$, and construct the alternative form of the covariance matrix:

$$\begin{pmatrix} \sigma_{p_1}^2 & \mathrm{cov}(p_1, p_0) \\ \mathrm{cov}(p_1, p_0) & \sigma_{p_0}^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_1^2} & \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_1 \partial p_0} \\ \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_1 \partial p_0} & \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_0^2} \end{pmatrix}^{-1}$$

4. Compare with the results of Exercise C.4. Would changing the step sizes alter the results?

---

## Summary

In this session you have seen how to relate the differential operator to the numerically more manageable difference operators (forward, backward and central). We did this for first order differentials and then extended it to higher orders. By following the exercises you should have convinced yourself that by basing the solution on central differences you obtain significantly faster convergence. Furthermore, as we have just seen, central differences are also more accurate for any particular order of solution.

By now you will appreciate that numerical analysis makes much use of expansions. Depending on your mathematics background you may or may not be very familiar with Taylor and Maclaurin series. Don't worry about this too much. In each session I try to give you a fairly rigorous derivation of the expressions that you need to use. My aim is to provide you with enough detail so that you can follow the algebra if you want. I won't test you on this, but you should have a working knowledge of the methodology and a firm understanding of the relative merits of different approaches.